Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 30 January 2002 (MN LaTeX style file v1.4)

Nonlinear Single-Armed Spiral Density Waves in Nearly Keplerian Disks

Erick Lee and Jeremy Goodman
Princeton University Observatory, Princeton, NJ, 08544–1001, U.S.A.

30 January 2002

ABSTRACT
Single-armed, stationary density waves can exist even in disks with only weak self-gravity, provided that the rotation curve is dominated by a central mass. Such waves could play a significant role in the transport of angular momentum. By variational methods, we derive nonlinear versions of the dispersion relation, angular momentum flux, and propagation velocity in the tight-winding limit. The pitch angle increases with amplitude until the tight-winding approximation breaks down. By other methods, we find a series of nonlinear logarithmic spirals which is exact in the limit of small disk mass and which extends to large pitch angle. These waves may be supported by low-mass protoplanetary disks, and perhaps by compact molecular disks in galactic nuclei.

Key words: accretion, accretion discs – galaxies: nuclei – hydrodynamics – Solar system: formation – waves

1 INTRODUCTION

The number of spiral arms in selfgravitating disks is influenced by the rotation curve. As observed by Lindblad, near-harmonic potentials and linearly rising rotation curves favor two-armed spirals: free-particle orbits are slowly precessing ellipses centered on the minimum of the potential; spiral waves can be built up from such orbits with relatively weak interactions needed to force them to precess at a common rate (cf. Binney & Tremaine (1987)). Since galactic potentials are often approximately harmonic at small radii, this construction may in part explain the prevalence of two-armed spirals and bars in disk galaxies.

The rotation curves of astrophysical disks smaller than galaxies are usually dominated by a central point-like mass: examples include accretion disks, protoplanetary disks, planetary rings, and masing molecular disks recently discovered in galactic nuclei (Section 4 and references therein). Free-particle orbits in such disks are slowly precessing keplerian ellipses. In this case, Lindblad's construction favors single-armed spirals, again with small angular pattern speed. Given the astrophysical importance of keplerian disks, it is unfortunate that single-armed spirals have received much less theoretical attention than double-armed ones. The visual fascination of beautiful spirals in galactic disks, and the absence of comparably well resolved images of keplerian disks, may partly explain the neglect.

The present paper is therefore dedicated to the propagation of single-armed spirals in nearly keplerian gaseous disks. We concentrate on the limit where the self-interaction of the disk is weak compared to its interaction with the central mass. This implies small or vanishing angular pattern speeds. The nonlinear regime is emphasized because it is gratifyingly tractable, and more importantly because one would like to know the maximum angular momentum flux that can be transmitted by density waves. We do not study the linear instabilities of single-armed waves (cf. Adams, Ruden, & Shu 1989; Shu et al. 1990; Heemskerk, Papaloizou, & Savonije 1992) or their excitation by companion masses (Yuan & Cassen 1985).

In Section 2, we adopt the familiar tight-winding approximation and extend it to nonlinear spirals. This has already been done using other methods, notably by Shu, Yuan, & Lissauer (1985); Shu, et al. (1985); and Borderies, Goldreich, & Tremaine (1985, 1986). Our mathematical approach, however, is adapted from the variational methods of Whitham (1974), which are particularly efficient for describing the propagation of a nonlinear wave train outside the region where it is excited, especially if dissipation is weak. In addition to the nonlinear dispersion relation and angular momentum flux, the variational approach provides the nonlinear analog of the group velocity, which as far as we know have not otherwise been derived for density waves (Appendix B). Also, previous work has mainly been concerned with nonlinear extensions of the long branch of the WKB dispersion relation, whereas we are primarily concerned with the short branch.

We show in Section 2 that for single-armed waves in keplerian potentials, self-gravity is best measured not by the familiar Q parameter (Toomre 1964), but by the ratio (2) of "Jeans length" (λ_J) to disk radius. In fact, almost all of our analyses pertain to the limit as $Q \to \infty$ at finite λ_J/r.

The results of Section 2 show that single-armed waves unwind with increasing nonlinearity. Therefore Section 3 presents nonlinear logarithmic spirals that do not depend upon the tight-winding approximation. The method of their construction assumes radial self-similarity, which restricts us to disks whose azimuthally averaged surface density scales with radius as $r^{-3/2}$. We believe, however, that many of the results are more generally valid. Solutions do not seem to exist unless $\lambda_J/r \lesssim 4$.

We summarize our results in Section 4 and discuss applications to astrophysical disks. In particular, we show that the minimum-mass solar nebula was probably in the regime considered here, namely large Q but moderate λ_J/r, so that stationary single-armed waves may have been important.

2 TIGHTLY-WRAPPED WAVES

In this section, we consider single-armed waves of small pitch angle. In Section 2.1, we summarize the linear theory, emphasizing those aspects that are peculiar to $m = 1$ waves in nearly keplerian potentials. These results provide background and guidance for the development of the nonlinear theory in Section 2.2, which itself has two parts: first we sketch a variational formalism for nonlinear wave trains in general, and then we apply this formalism to tightly-wrapped spiral density waves.

2.1 Linear Theory

The dispersion relation for tightly wrapped $m = 1$ spiral waves is

$$(\omega - \Omega)^2 = \kappa^2 - 2\pi G\Sigma k + c^2 k^2. \tag{1}$$

However it is excited, as the wave propagates inward, the pattern speed (ω) becomes negligible when compared to the local angular velocity of the disk (Ω). We study the stationary limit, $\omega \ll \Omega$, which is appropriate to waves far inside their corotation radius close to a central point mass. In this region, both κ^2 and Ω^2 are dominated by contributions from the central mass, but these large terms cancel one another to leading order in equation (1), leaving a residual due only to the disk.

To simplify the analysis we consider a razor thin disk and take a polytropic equation of state $P = K\Sigma^\gamma$. For a logarithmic spiral pattern, the density and temperature profiles should be power laws, $\Sigma \propto r^{-\beta-1}$ and $c^2 \propto r^{-\beta}$ where $0 < \beta < 1$. This requires $\gamma = (1+2\beta)/(1+\beta)$, so γ can range from 1 to 3/2. Now $\kappa^2 - \Omega^2$, the self-gravity term, and the pressure term in the dispersion relation all have the same scaling with r if the pitch angle $\cot^{-1}(kr) \approx 1/kr$ is constant. Introducing the dimensionless parameters

$$\sigma^2 \equiv \frac{c^2}{2\pi G\Sigma r} = \frac{\lambda_J}{2\pi r}, \tag{2}$$

and

$$g(\beta) \equiv \frac{(\kappa^2 - \Omega^2)r}{\pi G\Sigma} \quad \approx 2 \ + 0.7536\beta(1-\beta) - 2\sigma^2(1-\beta^2) \quad \text{if} \ \ 0 \le \beta \le 1. \tag{3}$$

we write the dispersion relation in dimensionless form

$$g - 2|k|r + 2\sigma^2(kr)^2 = 0. \tag{4}$$

Notice that $\kappa^2 - \Omega^2$ is independent of the central point mass. The term involving σ^2 in eq. (3) expresses the influence of pressure on the disk rotation curve; the rest is due to the disk's self-gravity. For small values of σ^2, the pitch angle $\cot^{-1}(kr) \approx \sigma^2$ for the physically relevant root of the dispersion relation (see Section 3.1), so the $\kappa^2 - \Omega^2$ term is negligible.

To see how the amplitude of the linear wave varies with radius we use the fact that wave angular momentum is conserved. The angular momentum flux due to spiral waves arises from gravitational and pressure torques, and from advective transport. The advective transport is proportional to $(1 - c^2|k|/2\pi G\Sigma)$ (Goldreich & Tremaine 1979) which vanishes to leading order in $1/kr$ for these waves. The flux due to gravitational and pressure torques can be calculated for tightly-wound patterns in the WKBJ approximation:

$$\Gamma_{WKB} = \mathrm{sign}(k)\frac{r\delta\Phi^2}{4G} = \mathrm{sign}(k)\frac{\pi^2 rG\delta\Sigma^2}{k^2} \tag{5}$$

Hence $d\Gamma_{WKB}/dr = 0$ implies $\delta\Sigma/\Sigma \propto r^{\beta-(1/2)}$, so that the fractional amplitude of the wave increases inward if $\beta < 1/2$.

The $m = 1$ waves have another special feature worth remarking upon here: in principle, they may exert a net force on the central mass. The component of the force along direction θ in polar coordinates is

$$\int_0^{2\pi}\int_0^{\infty} \frac{G\delta\Sigma(r,\theta')}{r^2}\cos(\theta-\theta')\,rdrd\theta.$$

If $\delta\Sigma(r,\theta') \propto \exp(im\theta')$, then the integral above can be nonzero for $m = 1$. In all cases of interest to us, however, $\delta\Sigma$ will be an oscillatory (wavelike) function of radius, so that the force is dominated by the endpoints of the radial integration. Thus, the displacement of the central mass cannot be included in a local analysis such as ours, whether linear or nonlinear, and we shall neglect it. For a careful global treatment of this issue in nonlinear as well as linear regimes, see Heemskerk, Papaloizou, & Savonije (1992). These authors find that modes capable of displacing the central mass tend to be evanescent in radius rather than wavelike.

2.2 Nonlinear Theory

Approximate solutions for a wide variety of nonlinear wave trains arising from mathematical physics can be found with remarkable efficiency by varational methods. Since these methods are not widely used in astrophysics, we digress briefly to summarize the main ideas. For a more complete exposition, see Whitham (1974).

The variational approach rests on the following assumptions:

(A.1) The fundamental equations of motion are derivable from an action principle, hence generally nondissipative.
(A.2) The wavelength and wave period of interest are small compared to the length scale and time scale of variations in the background.
(A.3) The waves are locally periodic in time and space, at least approximately.

Assumption (A.1) can be relaxed by tacking dissipative terms onto the equations of motion at a late stage in the procedure, but we will not do so here. Assumptions (A.2) and (A.3) are shared with linear WKBJ theory, but since the principle of superposition does not hold, one cannot synthesize a solitary wave—a highly localized wave packet—by linearly combining periodic wave trains. On the other hand, the third assumption permits gradual changes in the wavelength, wave period, and wave amplitude, provided that they occur on scales long compared to the wavelength and period.

With these assumptions, the following procedure accurately approximates nonlinear wave trains. The dynamical variable (u) depends in the first instance on one or more independent spatial variables ($x_1,\ldots,x_n$) and on time (t); u might in fact represent multiple dependent variables. By assumption (A.1), u obeys an action principle:

$$\delta\int L(u,\partial_t u,\partial_1 u,\ldots,\partial_n u;x_1,\ldots,x_n,t)\mathcal{J}(x)d^n x dt = 0. \tag{6}$$

Here $\mathcal{J}(x)$ is just the Jacobian of the transformation from $(x_1,\ldots,x_n)$ to cartesian coordinates. We have indicated that L depends on first spatial derivatives of u, but in fact it may depend on spatial derivatives of any order, as is the case for self-gravitating density waves. For solutions of the action principle obeying assumption (A.3), u can be written as a periodic function of a single phase (Ψ):

$$u = U(\Psi) = U(\Psi+2\pi), \qquad \Psi = \Psi(x_1,\ldots,x_n,t). \tag{7}$$

The frequency and wavenumber are defined as derivatives of Ψ:

$$\omega \equiv -\partial_t\Psi, \qquad k_i \equiv \partial_i\Psi \equiv \frac{\partial\Psi}{\partial x_i}. \tag{8}$$

Hence $\partial_t u \to -\omega U'$ and $\partial_i u \to k_i U'$. The critical approximation is now to *average over the phase* Ψ *while pretending that* $\omega, k_1,\ldots,k_n$ *and* $t, x_1,\ldots x_n$ *are constant*:

$$\bar{L}(U,U',\omega,k_1,\ldots;x_1,\ldots,t) \equiv \int_0^{2\pi} L(U,-\omega U',k_1U',\ldots;x_1,\ldots,t)\mathcal{J}(x)\frac{d\Psi}{2\pi}. \tag{9}$$

The action principle (6) reduces to

$$\delta\int \bar{L}(U,U';\omega,k_1,\ldots;x_1,\ldots,t)\mathcal{J}(x)d^n x dt = 0, \tag{10}$$

where δ now indicates variation of $U(\Psi)$ and $\Psi(x_1,\ldots,x_n,t)$.

Stationarity of the action with respect to Ψ, which enters only through its derivatives (8), implies the conservation law

$$\frac{\partial}{\partial t}\left(\mathcal{J}\frac{\partial\bar{L}}{\partial\omega}\right)-\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left(\mathcal{J}\frac{\partial\bar{L}}{\partial k_i}\right)=0. \tag{11}$$

The quantity $\partial\bar{L}/\partial\omega$ is the density of wave action, and $-\partial\bar{L}/\partial k_i$ is the $i^{\rm th}$ component of the corresponding flux.

Stationarity of the action with respect to U yields both the nonlinear dispersion relation and also equations for the functional form of $U(\Psi)$ [cf. Whitham 1974].

To apply this formalism to spiral density waves, the radial wavelength must be small compared to the radius in the disk (r), since the surface density and other background properties vary significantly over radial distances comparable to r. For single-armed spirals, a short radial wavelength implies that the wave is tightly wrapped. So we may as well take advantage of tight winding to simplify our formulae.

For dynamical variables, we take the radial and angular displacements of a fluid particle from a circular orbit:

$$\delta r\equiv r_{\rm p}(t),\qquad \delta\theta=\theta_{\rm p}(t)-\Omega(a)t-\theta_{\rm p}(0), \tag{12}$$

where $[r_{\rm p}(t),\theta_{\rm p}(t)]$ are the polar coordinates of the particle at time t. Here a and $\theta_p(0)$ are labels that follow the particle, but we will soon interpret a as the semimajor axis of a keplerian orbit that closely approximates the particle's trajectory. $\Omega(a)$ is the mean angular velocity of that trajectory, which is not exactly the same as the mean motion of a keplerian orbit of semimajor axis a, because the pressure and self-gravity of the disk modify its rotation curve. For independent variables we take (a,θ,t). In tightly-wound waves where the streamlines do not actually cross, $\delta r/a\ll 1$ but $\partial\delta r/\partial a$ may approach unity. Therefore in constructing the lagrangian, we keep only the lowest important order in $\delta r/a$ but all orders in $\partial\delta r/\partial a$.

In the absence of any collective effects, the lagrangian and the action would be those of a collection of free particles:

$$I_{\rm free}=\int_{t_1}^{t_2}\int_0^{2\pi}\int_0^{\infty}L_{\rm free}(\delta r,\delta\dot{r},a)\,a da d\theta dt,\qquad L_{\rm free}=\frac{1}{2}\left[\left(\frac{d\delta r}{dt}\right)^2-\kappa^2(\delta r)^2\right]\Sigma_0(a), \tag{13}$$

where

$$\kappa^2\equiv\frac{1}{2a^3}\frac{d}{da}\left(a^4\Omega^2\right) \tag{14}$$

is the epicyclic frequency, and $\Sigma_0(a)$ is the surface density that fluid elements would have in the unperturbed axisymmetric disk. The corresponding Euler-Lagrange equation for each fluid element is $\delta\ddot{r}=-\kappa^2 r$.

To represent the self-interaction of the disk, we subtract from $I_{\rm free}$ the time integral of two potential-energy terms. The first of these is the thermodynamic internal energy, $W_{\rm int}$. Absent dissipation, the two-dimensional pressure and density of a given fluid element scale as $P\propto\Sigma^{\gamma}$, and the internal energy per unit mass is $\gamma P/(\gamma-1)\Sigma$. Let $P_0(a)$ be the 2D pressure that a fluid element would have in a circular orbit. Its actual surface density is

$$\Sigma(a,\theta,t)=\Sigma_0(a)\left(1+\frac{\partial\delta r}{\partial a}\right)^{-1} \tag{15}$$

in the tight-winding approximation. It follows that the total internal energy of the gas is

$$W_{\rm int}=\int\!\!\int\frac{c_0^2(a)}{\gamma(\gamma-1)}\left[\left(1+\frac{\partial\delta r}{\partial a}\right)^{-(\gamma-1)}-1\right]\Sigma_0(a)a da d\theta, \tag{16}$$

where $c_0^2(a)\equiv\gamma P_0/\Sigma_0$ is the square of the unperturbed sound speed.

The gravitational self-energy of a tightly-wound wave is accurately expressed by the "thin-wire" approximation. On the scale of the radial wavelength, curves of constant surface density are approximately straight lines. The gravitational interaction energy of two lines of mass per unit length (μ_1,μ_2) and separation s_{12} is $G\mu_1\mu_2\ln|s_{12}|$ per unit length. We put $\mu_1=\Sigma_0(a)da_1$, $\mu_2=\Sigma_0(a)da_2$, where $a\equiv(a_1+a_2)/2$, since $\Sigma_0(a)$ varies only on the scale of a. And since the pitch angle of the streamlines is small, $s_{12}=|a_1+\delta r_1-(a_2+\delta r_2)|$, where $\delta r_i\equiv r(a_i,\theta,t)$ and $\theta\equiv(\theta_1+\theta_2)/2$. Therefore,

$$W_{\rm grav}=G\int_0^{2\pi}d\theta\int_0^{\infty}a\Sigma_0^2(a)da\int_{-2a}^{2a}d(a_1-a_2)\ \ln\left|1+\frac{\delta r_1-\delta r_2}{a_1-a_2}\right|. \tag{17}$$

We have replaced s_{12} by $s_{12}/(a_1-a_2)$ inside the logarithm; this changes $W_{\rm grav}$ by a term independent of δr, which has no effect on the equations of motion. The integration over (a_1-a_2) is also insensitive to its limits, since the dominant contribution comes from $|a_1 - a_2| \ll a$ in a tightly-wound wave. The above expression for $W_{\rm grav}$ describes local self-gravity but not the long-range contribution of the unperturbed axisymmetric disk surface density to the rotation curve. The latter is included in the mean angular velocity $\Omega(a)$ and in the epicyclic frequency (14) derived from it.

For single-armed wavelike solutions, δr and $\delta\theta$ depend on (a, θ, t) *via* the phase variable

$$\Psi \equiv \theta - \phi(a) - \omega t. \tag{18}$$

In principle the wavenumber has two components, but since the angular component $k_\theta \equiv \partial\Psi/\partial\theta$ is always unity, we write

$$k \equiv \frac{\partial\Psi}{\partial a} = -\phi'(a) \tag{19}$$

for the radial component.

Our next step is to average the full lagrangian over one period in Ψ. This task is simplified by the fact that we are interested in situations where the interior mass of the disk is small compared to that of the central object (M). We introduce a small parameter

$$\eta \equiv \frac{2\pi\bar{a}^2\Sigma_0(\bar{a})}{M} \ll 1, \tag{20}$$

where $\bar{a}$ is typical of the radii of interest. The self-interaction terms $W_{\rm int}$ and $W_{\rm grav}$ are formally of order η compared to the free action (13). To leading order in η, the trajectory of a fluid element must therefore be a solution of the Euler-Lagrange equation derived from $L_{\rm free}$ alone. Such a solutions is sinusoidal in t, and since it must depend on t through Ψ, it can be written in the form

$$\delta r(a, \theta) = a e(a) \cos\Psi, \tag{21}$$

where e is a dimensionless measure of the amplitude. For $\eta = 0$, the only forces would be those of an exactly keplerian potential; in that case, $\omega = 0$, and e and ϕ could be chosen as arbitrary constants independently for each fluid element. For $\eta \neq 0$ but small, e and ϕ must be smooth functions of a because the self-interaction terms would diverge if neighboring streamlines were to intersect. By substituting the form (21) into the lagrangian and applying the phase-averaged action principle, one finds relations between e, ω, and k to leading order in η. At higher orders in η, there are corrections to these relations and to the functional form (21).

Note that the small parameter η is independent of the degree of nonlinearity of the wave. The latter is measured in the tight-winding approximation by the streamline-crossing parameter

$$q_0 = |k| a e \tag{22}$$

Streamlines intersect if and only if $q_0 \geq 1$, as can be seen by differentiating eq. (21) with respect to a assuming $d\ln e/da = O(1/a) \ll k$, and then referring to eq. (15). We will assume that $q_0 < 1$, but not necessarily $q_0 \ll 1$, since otherwise dissipation is inevitable. Hence since $ka \gg 1$ is necessary for the tightwinding approximation, we must have $e \ll 1$.

The time derivative in eq. (13) is the total time derivative following a fluid element,

$$\frac{d\delta r}{dt} \approx \frac{\partial \delta r}{\partial t} + \Omega(a)\frac{\partial\delta r}{\partial\theta}.$$

The angular velocity $\dot{\theta} = \Omega + O(e)$, but we ignore the $O(e)$ correction when it multiplies angular derivatives, since these are small compared to radial derivatives. Hence the phase average of $L_{\rm free}$ is

$$\int_0^{2\pi} L_{\rm free}\frac{d\Psi}{2\pi} = \frac{\Sigma_0(a)}{4}\left[(\Omega - \omega)^2 - \kappa^2\right] e^2(a). \tag{23}$$

The phase averages of the internal and gravitational energies are derived in Appendix A, and are expressed in terms of the following integral functions:

$$\mathcal{U}(q_0) \equiv \frac{4}{\gamma(\gamma-1)} q_0^{-2} \int_0^{2\pi} \frac{d\Psi}{2\pi}\left[(1 - q_0\sin\Psi)^{-(\gamma-1)} - 1\right], \tag{24}$$

and

$$\mathcal{W}(q_0) \;\equiv\; -\frac{4}{\pi|q_0|}\int_{-\infty}^{\infty} dy \ln\left|\frac{1}{2} + \frac{1}{2}\sqrt{1 - \left(\frac{\sin(q_0 y)}{y}\right)^2}\right|. \tag{25}$$

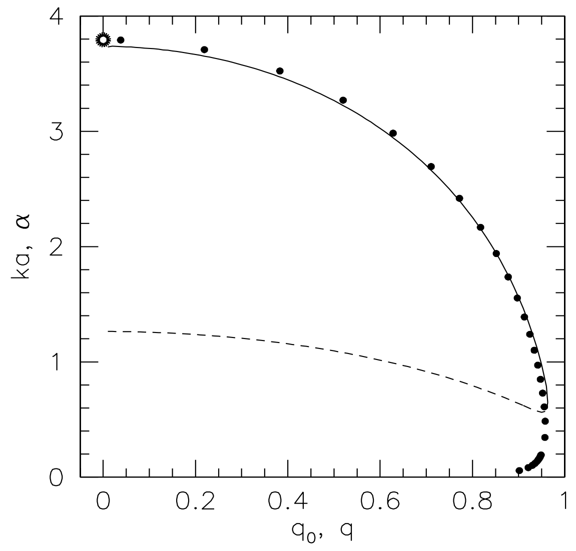

Figure 1. Nonlinear dispersion relation for stationary single-armed waves at $\sigma^2 = 0.2$, $\beta = 1/2$, and $\gamma = 4/3$. *Abscissa*: streamline-crossing parameter q_0 (eq. [22]) or q (eq. [50]). *Ordinate*: dimensionless radial wavenumber ka or α, equal to cotangent of pitch angle. *Solid curve:* Short branch in tight-winding theory (ka vs. q_0). *Dashed curve:* Unphysical long branch. *Filled Points:* Logarithmic spirals (α vs. q). *Hollow Point:* Linear theory (Section 3.1)

These functions are normalized so $\mathcal{U} = 1 + \mathcal{O}(q_0^2)$ and $\mathcal{W} = 1 + \mathcal{O}(q_0^2)$. Both increase monotonically with q_0 and both are finite but nondifferentiable at $q_0 = 1$: $\mathcal{U}'(q_0) \propto (1-q_0)^{1/2-\gamma}$ as $q_0 \to 1$, while $\mathcal{W}'(q_0) \propto \ln(1-q_0)$.

The phase-averaged action is therefore

$$I = \int\int \bar{L}(e, w, k, a) 2\pi a da dt. \tag{26}$$

It is convenient to divide $\bar{L}$ into a part $\bar{L}_0$ that is independent of ω and a residual:

$$\begin{aligned} \bar{L}(e,\omega,k,a) &= \bar{L}_0(e,k,a) + \frac{\Sigma_0(a)}{4}\omega[\omega - 2\Omega(a)]a^2e^2, \\ \bar{L}_0(e,k,a) &= -\frac{\pi G\Sigma_0^2 a}{2}\left[\frac{1}{2}ge^2 - |q_0|\mathcal{W}(q_0)e + \sigma^2 q_0^2\mathcal{U}(q_0)\right] \end{aligned} \tag{27}$$

The averaged action (26) must be stationary with respect to e; imposing $\partial\bar{L}/\partial e = 0$ yields

$$ge^2 - e\,\mathrm{sign}(k)\frac{d}{dq_0}[q_0^2\mathcal{W}(q_0)] + \sigma^2 q_0\frac{d}{dq_0}[q_0^2\mathcal{U}(q_0)] - \omega(\omega - 2\Omega)\frac{ae^2}{\pi G\Sigma_0} = 0. \tag{28}$$

In the limit $e \ll 1$ at fixed ka, every term in this expression is proportional to e^2, and we recover the linear dispersion relation (1) or (4), since $\mathcal{U}(0) = \mathcal{W}(0) = 1$ and $\mathcal{U}'(0) = \mathcal{W}'(0) = 1$. In general, however, the dispersion relation relates the amplitude of the wave to its wavelength and frequency. If we fix one of these three quantities, then the other two are constrained to a curve. In this paper, we are most interested in $\omega = 0$. Figure 1 displays some representative dispersion curves for stationary waves. We have used q_0 rather than e as a measure of wave amplitude because it more directly controls the importance of pressure and gravitational forces.

The Euler-Lagrange equation for Ψ is (11), with the jacobian $\mathcal{J} = 2\pi a$ in this case. Hence the wave action density (action per unit physical area) is

$$\rho \equiv \frac{\partial\bar{L}}{\partial\omega} = (\omega - \Omega)\frac{\Sigma_0 a^2 e^2}{2}. \tag{29}$$

This can also be interpreted as the angular momentum density of the wave. The density is negative for stationary waves ($\omega \to 0$). The corresponding radial flux is

$$f \equiv -\frac{\partial\bar{L}}{\partial k} = \frac{\pi G\Sigma_0^2 a^2 e}{2}\frac{\partial}{\partial q_0}\left[\sigma^2 q_0^2\mathcal{U}(q_0) - e|q_0|\mathcal{W}(q_0)\right]$$

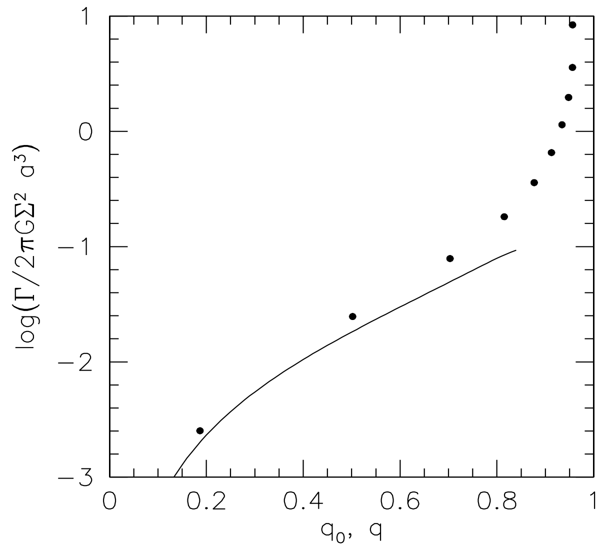

Figure 2. Wave torque (azimuthally integrated radial angular momentum flux) for $\sigma^2 = 0.2$, $\beta = 1/2$, $\gamma = 4/3$. *Solid curve:* Tight-winding theory; $\Gamma = 2\pi af$, f from eq. (30) with $\omega = 0$. *Points:* Torque of logarithmic spirals (Section 3.3 and Appendix C).

$$= \frac{\pi G\Sigma_0^2 a^2 e^2}{2}\left[\mathrm{sign}(k)\mathcal{W}(q_0) - \frac{g}{ka}\right] + \omega(\omega - 2\Omega)\frac{\Sigma_0 a^2 e^2}{k}. \tag{30}$$

In the second line, we have used the dispersion relation (28) to eliminate the derivatives of $\mathcal{U}$ and $\mathcal{W}$. Strictly, the term g/ka is negligible in the tightly-wound limit. From eqs. (15) and (21), the fractional surface density fluctuation becomes $\delta\Sigma/\Sigma = q_0 = kae$ for $q_0 \ll 1$. Therefore in the linear (and stationary and tightly-wound) limit, $2\pi af$ reduces to Γ_{WKB} (eq. [5]).

For stationary waves, $\Gamma \equiv 2\pi af$ is independent of radius. This quantity contains the grouping of disk parameters $\Sigma^2 a^3$, which scales with radius as $a^{1-2\beta}$. Therefore if $\beta = 1/2$, a stationary wave occupies a single point (q_0, ka) on a dispersion curve such as those in Fig. 1 at all radii. If $\beta \neq 1/2$, a radially-propagating wave train moves along the dispersion curve to compensate for the changes in the disk. This defines q_0 and e as functions of a, with the angular momentum flux Γ as a parameter. In particular, if $\beta < 1/2$, tightly-wound trailing waves tend to become more nonlinear as they propagate inward, because $\Gamma/(\Sigma^2 a^3)$ tends to increase with q_0, as is shown for $\sigma^2 = 0.2$ in Figure 2. Unfortunately, as Figure 1 demonstrates, the pitch angle $\cot^{-1}(ka)$ also increases with q_0 until the tight-winding approximation can no longer be trusted. This generally happens well before the point of streamline crossing is reached ($q_0 = 1$).

The stationary dispersion relation (28) can be regarded as a quadratic equation in e with coefficients in q_0. The two solutions are analogous to the short and long branches of linear theory. Since $k = q_0/ae$, the smaller solution for e is the short branch. The discriminant is

$$D(q_0) = \left[\frac{d}{dq_0}(q_0^2\mathcal{W})\right]^2 - 4g\sigma^2 q_0 \frac{d}{dq_0}(q_0^2\mathcal{U}) \tag{31}$$

The eccentricity e is real only if $D(q_0) > 0$. Since the derivative of $\mathcal{U}$ diverges faster than the derivative of $\mathcal{W}$ as $q_0 \to 1$, there is some $q_0 < 1$ beyond which no real solutions for the eccentricity exist. At this critical q_0, the short and long branches join. These conclusions are academic, however, since the join occurs outside the tightly-wound regime. In fact, the long branch does not actually exist for stationary waves (Section 3).

We have tacitly assumed that short-branch stationary trailing waves always propagate inwards. To verify this assumption for our nonlinear wave trains, we must investigate the nonlinear generalization of the group velocity. An obvious candidate is the ratio f/ρ of wave flux to wave density, which reduces to the radial group velocity in the linear limit. Nonlinearly, however, there are actually two distinct characteristic velocities at which information about the wave propagates; neither coincides with f/ρ or v_{group} except as $q_0 \to 0$. As shown by Whitham (1974), this is a generic property of nonlinear wave trains. The characteristic velocities in the present case are derived in Appendix B and exemplified in Figure 3. Both characteristic velocities are negative (inward) for trailing waves, at least as long as the wave remains tightly wound.

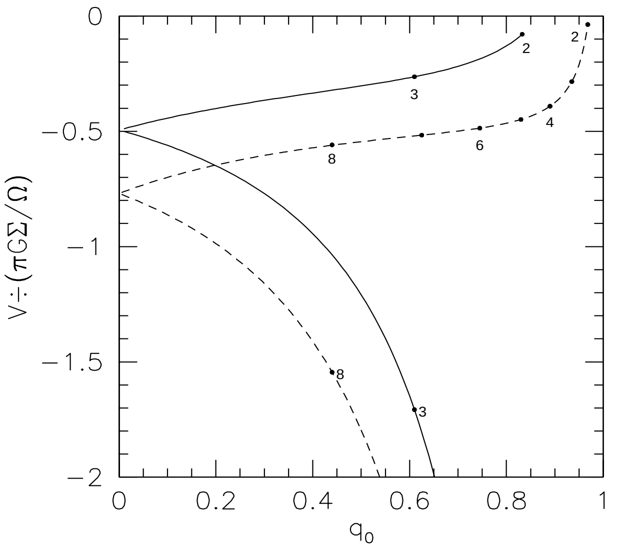

Figure 3. The two characteristic velocities (radial propagation velocities) of stationary, trailing, single-armed spirals in the tight-winding approximation, for $\beta = 1/2$, $\gamma = 4/3$, $\sigma^2 = 0.2$ (*solid curves*) and $\sigma^2 = 0.1$ (*dashed*). The two velocities converge to the linear group velocity as the streamline-crossing parameter $q_0 \to 0$. Labeled points mark radial wavenumber ka along each sequence. Smaller σ^2 makes for tighter winding (larger ka) at a given q_0.

3 EXACT SELF-SIMILAR SPIRALS

The nonlinear wave-action formalism of Section 2 is flexible but assumes tightly-wrapped waves, $|kr| \gg 1$. In this section, we sacrifice generality for accuracy and find solutions in the form of exact logarithmic spirals of arbitrary pitch angle.

The surface density is

$$\Sigma(r,\theta) = r^{-3/2} S(\theta + \alpha \ln r), \tag{32}$$

where S is periodic $[S(\psi+2\pi) = S(\psi)]$ but may be nonsinusoidal. The cotangent of the pitch angle of the spiral, α, corresponds to the previous ka. As suggested by the discussion of flux conservation in section Section 2, the surface density must vary as $r^{-3/2}$ on average if the spiral is to have the same strength $\delta\Sigma/\Sigma$ at all radii. Self-similarity gives the same requirements as in Section 2; that the sound speed vary as $r^{-1/2}$, and the adiabatic index must be $\gamma = 4/3$.

Syer & Tremaine (1996) have studied self-similar nonaxisymmetric power-law disks. Their solutions omit the central point mass and assume a constant ratio of sound speed to orbital speed (at a given phase in the spiral), as is required by strict self-similarity. In our case, this ratio is not constant: it varies as $r^{(1-\beta)/2}$ in general, and as $r^{1/4}$ in this section. Nevertheless our models are self-similar when viewed in the limit $c/v_{\rm orb} \to 0$, or equivalently $r^2\Sigma(r)/M \to 0$, and it is only in this limit that they are exact. We imagine that we are studying waves very close to the central mass.

3.1 Linear theory

Much insight into the nonlinear dispersion relation can be obtained from the linear limit, where the surface density function (32) reduces to

$$\Sigma(r,\theta) = S_0 r^{-3/2}[1 + \epsilon\cos(\theta + \alpha\ln r)], \qquad \Sigma_0(r) + \epsilon\Sigma_1(r,\theta), \tag{33}$$

with S_0 a constant and $\epsilon \ll 1$. The potential corresponding to (33) is [Kalnajs 1971]

$$\Phi(r,\theta) = -2\pi G S_0 r^{-1/2}\left[K(0,0) + \epsilon K(\alpha,1)\cos(\theta+\alpha\ln r)\right], \qquad \equiv \Phi_0(r) + \epsilon\Phi_1(r,\theta), \tag{34}$$

where

$$K(\alpha,m) = \frac{1}{2}\,\Gamma\left(\frac{2m+1+2i\alpha}{4}\right)\Gamma\left(\frac{2m+1-2i\alpha}{4}\right)\Big/\left[\Gamma\left(\frac{2m+3+2i\alpha}{4}\right)\Gamma\left(\frac{2m+3-2i\alpha}{4}\right)\right]. \tag{35}$$

Linearization of the standard inviscid equations of motion yields the following dispersion relation:

$$\kappa^2 - (m\Omega - \omega)^2 - \frac{2\pi G\Sigma_0(r)}{r}K(\alpha, m)\left(\alpha^2 + \frac{9}{4}\right) + \frac{c^2(r)}{r^2}\left(\alpha^2 + \frac{9}{4}\right) = 0. \tag{36}$$

In the tight-winding limit $\alpha \gg m \geq 1$, $K(\alpha, m) \approx |\alpha|^{-1}$, so that eq. (1) is recovered. Allowing for the effects of the effects of pressure on the equilibrium rotation curve,

$$\frac{(\kappa^2 - \Omega^2)r}{\pi G\Sigma_0(r)} = \frac{1}{2}[K(0,0) - 3\sigma^2], \tag{37}$$

(cf. eq. [3]), so that the dispersion relation for $m = 1$ and $\omega = 0$ can be put in dimensionless form:

$$K(0,0) - 3\sigma^2 + (4\alpha^2 + 9)[\sigma^2 - K(\alpha, 1)] = 0. \tag{38}$$

For any choice of σ^2, one might expect two positive roots for α, corresponding to the intersection of the long and short branches with $\omega = 0$.* However, equation (38) has at most one root. Whereas the tight-winding equivalent eq. (4) is positive as $kr \to 0$, the left hand side of eq. (38) remains negative as $\alpha \to 0$ because of the term involving $K(\alpha, 1)$. Self-gravity is apparently too strong to allow the long-branch wave to be stationary (see below).

The dispersion relation has no real roots at all if $\sigma^2 \gtrsim 0.64136$. At least for $\beta = 1/2$ and $\gamma = 4/3$, the maximum Jean's length that permits these single-armed stationary waves to exist is

$$\left(\frac{\lambda_J}{r}\right)_{\max} \equiv 2\pi(\sigma^2)_{\max} = 4.030. \tag{39}$$

It is clear that the real root for α, where it exists, belongs to the short branch. Since the long-branch wave depends primarily on self-gravity, its wavenumber should be independent of $\sigma^2 \propto c^2$ as $c^2 \to 0$, yet eq. (38) dictates that $\alpha \approx \sigma^{-2}$ as $\sigma^2 \to 0$. [$K(\alpha, m) \approx |\alpha|^{-1/2}$ for $\alpha \gg 1$.] However, the wave is "short" only in a relative sense. For $m \neq 1$ and except near resonances, the short wave normally has a wavelength comparable to the disk thickness or smaller. In this case, the typical wavelength is $\sim \lambda_J$, which is large compared to the disk thickness in the limit that $M_{\rm disk} \ll M_{\rm star}$ (or equivalently $Q \to \infty$) at fixed σ^2. For our purposes, the thin-disk approximation is almost always appropriate.

To recap, *there is no stationary long wave.* However, there exist long waves with prograde pattern speeds. Equation (37) shows that $\kappa^2 - \Omega^2 \propto r^{-5/2}$, whereas $\Omega \propto r^{-3/2}$. To allow a long wave at small α, the dispersion relation (36) requires $\omega > 0$ and $\omega \propto r^{-1}$. As $r \to 0$, ω is asymptotically negligible compared to $\Omega(r)$ but asymptotically infinite compared to the orbital frequency of any perturbing mass that might launch waves inward.

3.2 Nonlinear dispersion relation

In the limit that the self-gravity and pressure of the disk are very weak, every fluid element must follow an orbit compatible with the dominant potential of the central mass. Hence the streamlines must be keplerian ellipses, although their eccentricity need not be small. The pressure and self-gravity of the disk control the choice of the eccentricity and spirality (e, α) in a manner described by the dispersion relation. Since our problem is no longer local in radius, we cannot use the elegant wave-action methods exploited in Section 2 and instead must look elsewhere for a physical condition to provide the dispersion relation.

Note first of all that the net torque per unit mass on each streamline must vanish, since the orbits are closed † But this condition does not provide the dispersion relation, because it can be shown that *the net torques on a streamline due to pressure and gravity vanish separately for all choices of e and α*. This is a property of logarithmic spirals for which the surface density varies as $r^{-3/2}$ and the sound speed as $r^{-1/2}$ along a wave crest. It is closely related to the fact that the angular momentum flux carried by such spirals is independent of radius [cf. Syer & Tremaine (1996) and our Appendix B].

The actual condition from which we derive the dispersion relation is that the streamlines should not precess. Streamlines become free-particle trajectories if one incorporates the pressure and self-gravity of the disk into an effective potential,

$$V(r, \theta) = -\frac{GM}{r} + \Phi(r, \theta) + H(r, \theta), \tag{40}$$

where the first term on the right is due to the central mass, the second term is the gravitational potential of the disk, and the third term is the enthalpy,

$$H \equiv \frac{c^2}{\gamma - 1}, \tag{41}$$

* Lynden-Bell & Ostriker (1967)'s Antispiral Theorem requires that all roots occur in pairs $\pm\alpha$: every trailing spiral ($\alpha > 0$) has a leading counterpart ($\alpha < 0$). We count only positive roots.

† Shocks and other forms of dissipation could allow the gas to absorb negative angular momentum from the waves, and then there would be a nonzero torque balanced by accretion.

which varies with the surface density as $\Sigma^{\gamma-1}$. (The adiabatic exponent appears in symbolic form for the sake of clarity, but $\gamma \to 4/3$.) The accelerations due to pressure are $-\nabla H$.

Let the time average of V along a keplerian ellipse be $\overline{V}(a, e, \phi)$, where ϕ is the periapse angle. Then by standard methods of secular perturbation theory, the condition that the orbit not precess is (cf. Brouwer & Clemence 1961)

$$\frac{\partial \overline{V}}{\partial e}(a, e, \phi) \equiv 0, \tag{42}$$

to which the keplerian part does not contribute.

It is convenient to introduce the eccentric anomaly E and the true anomaly ψ. The relation between the two anomalies and the time in a keplerian orbit is (cf. Brouwer & Clemence 1961)

$$E - e\sin E = \psi \equiv n(t - t_{\rm peri}). \tag{43}$$

Here $n = 2\pi(a^3/GM)^{1/2}$ is the mean motion, and $t_{\rm peri}$ is the time of pericenter. If ϕ is the azimuth of pericenter, the position in polar coordinates (r, θ) is

$$\begin{aligned} r &= a(1 - e\cos E), \\ r\cos[\theta - \phi(a)] &= a(\cos E - e), \\ r\sin[\theta - \phi(a)] &= a\sqrt{1-e^2}\sin E. \end{aligned} \tag{44}$$

Because of self-similarity, the disk potential and enthalpy take the forms

$$\Phi(a, E) = a^{-1/2} f(E), \qquad H(a, E) = a^{-1/2} h(E), \tag{45}$$

where $f(E)$ and $h(E)$ are periodic functions. It is sufficient to calculate f and h along a single streamline.

One can calculate $f(E)$ as a double integral over all other streamlines:

$$\begin{aligned} f(E) &= -a^{1/2} G \int_{a'=0}^{\infty} da'\, a'\Sigma_0(a') \int_0^{2\pi} dE' \frac{1 - e\cos E'}{R}, \\ R^2 &\equiv a^2(1 - e\cos E)^2 + a'^2(1 - e\cos E')^2 - 2aa' \left\{ \left[(\cos E' - e)(\cos E - e) + (1 - e^2)\sin E \sin E' \right] \cos(\alpha \ln a') \right. \\ & \qquad \left. - \left[(\cos E' - e')\sin E - (\cos E - e)\sin E' \right] \sqrt{1-e^2}\sin(\alpha \ln a') \right\}. \end{aligned} \tag{46}$$

Here R is the distance between the points corresponding to E and E' on the streamlines a and a', respectively. The factor $(1 - e\cos E')$ is proportional to the time spent in the interval dE', and $\Sigma_0(a)$ is a lagrangian surface density defined so that the mass between streamlines a and $a + da$ as $2\pi\Sigma_0(a)ada = 2\pi\Sigma_0(1)a^{-1/2}da$.

A more efficient way to calculate the gravitational potential is to decompose the surface density $\Sigma(r, \theta)$ into a sum of azimuthal Fourier harmonics, each itself a logarithmic spiral, and then to use Kalnajs' formula to obtain the corresponding harmonics of $\Phi(r, \theta)$:

$$\Sigma(r, \theta) = r^{-3/2} \sum_{m=0}^{\infty} \hat{\Sigma}(m) \cos m(\theta + \alpha \ln r), \qquad \Phi(r, \theta) = -2\pi G r^{-1/2} \sum_{m=0}^{\infty} K(m\alpha, m)\hat{\Sigma}(m) \cos m(\theta + \alpha \ln r). \tag{47}$$

Since $\Sigma(r, \theta)$ has the form (32), the angular Fourier transform needed to find the coefficients $\hat{\Sigma}(m)$ need be performed only at one radius. But we did not use this method except as a check.

The enthalpy and surface density are easily calculated in streamline coordinates. The transformation (44) from $(a, \psi) \to (r, \theta)$, when e is fixed and $\phi(a) = -\alpha \ln a$, has the Jacobian

$$\mathcal{J} \equiv \frac{r dr d\theta}{a da d\psi} = \sqrt{1-e^2} + \alpha e \sin E. \tag{48}$$

Since

$$\frac{d\psi}{dE} = 1 - e\cos E = \frac{r}{a},$$

we also have $dr d\theta = \mathcal{J} da dE$. If the wave is an adiabatic deformation of a circular disk, the gas is uniformly distributed with respect to orbital phas ψ, so that $\Sigma r dr d\theta = \Sigma_0 a da d\psi$, whence

$$\Sigma(a, E) = \Sigma_0(a)\mathcal{J}^{-1}, \qquad H(a, E) = \sigma^2 \cdot 2\pi G \Sigma_0(a) a \mathcal{J}^{1-\gamma}. \tag{49}$$

Now $\mathcal{J}$ has zeros and Σ has poles unless the streamline-crossing parameter

$$q \equiv \alpha e/\sqrt{1-e^2} \tag{50}$$

is less than unity. So q replaces the quantity $q_0 \equiv kae$ of the tightly-wrapped theory.

The machinery above allows one to calculate the averages of Φ, H, and hence V around the actual streamlines. For example,

$$\overline{\Phi} = \int_0^{2\pi} \Phi \frac{d\psi}{2\pi} = \int_0^{2\pi} \frac{dE}{2\pi}(1-e\cos E)\Phi(a,E). \tag{51}$$

But the streamline on which the average above is taken has the same shape as all of the others, whereas to compute the variation (42), one needs to vary the eccentricity of the "test" streamline while holding all the rest—the "field" streamlines—fixed. The potentials Φ, H, and V are regarded as fixed functions of the Eulerian coordinates $(a_{\rm f}, E_{\rm f})$ based on the field streamlines, all of which have eccentricity $e_{\rm f}$ and spirality $\alpha = -d\phi_{\rm f}/d\ln a_f$: the subscripts "f" and "t" indicate field and test quantities, respectively. A calculation yields

$$\left(\frac{\partial a_{\rm f}}{\partial e_{\rm t}}\right)_{a_{\rm t},E_{\rm t},\phi_{\rm t}} = -\frac{a(e+\cos E)}{\mathcal{J}\sqrt{1-e^2}}, \qquad \left(\frac{\partial E_{\rm f}}{\partial e_{\rm t}}\right)_{a_{\rm t},E_{\rm t},\phi_{\rm t}} = \frac{\sin E - \alpha\sqrt{1-e^2}\cos E}{\mathcal{J}\sqrt{1-e^2}}, \tag{52}$$

in which $(a_{\rm t,f}, E_{\rm t,f}, \phi_{\rm t,f}) \to (a, E, \phi(a))$ after the differentiation. Hence

$$\frac{\partial\overline{\Phi}}{\partial e} = a^{-1/2}\int_0^{2\pi} \frac{dE}{2\pi}\left[\frac{1}{2}(e+\cos E)\,f(E) + \left(\sin E - \alpha\sqrt{1-e^2}\cos E\right)\frac{df}{dE}\right]\frac{1-e\cos E}{\mathcal{J}\sqrt{1-e^2}}. \tag{53}$$

The precession rate due to pressure is given by an expression identical to the above except that $\overline{H}$ & h replace $\overline{\Phi}$ &f .

The results for $\sigma^2 = 0.2$ are shown by the points in Fig. 1. We have checked the results for $e \ll 1$ against the linear theory of Section 3.1: $\alpha \approx 3.794$ (linear); $\alpha \approx 3.792$ at $e = 0.01$ (nonlinear). The dispersion relation predicted by the tight-winding theory of Section 2 is also included in Figure 1 using $q_0 \equiv kae$ rather than q as the abscissa. Both the present self-similar solutions and the tight-winding theory of Section 2 predict that the spirals unwind with increasing nonlinearity ($d\alpha/dq < 0$). But the long branch predicted by the tight-winding approximation does not exist. Notice that like the tight-winding dispersion relation, the self-similar sequence reaches a maximum q. The linear analysis of Section 3.1 shows, however, that the self-similar sequence cannot return to $q = 0 = e$ at nonzero wavenumber α. Therefore it must terminate in a point at $\alpha = 0$, $q > 0$, and $e = 1$. The last (lowest) self-similar model shown has $e = 0.998$, $q = 0.9016$, and $\alpha = 0.057$, which corresponds to a pitch angle of $86.7°$. Our numerical methods are unable to follow the sequence beyond this point.

3.3 Angular momentum flux

There is more than one way to define the angular momentum flux of a spiral wave. Lynden-Bell & Kalnajs (1972, henceforth LBK) imagine dividing the disk in two by a cylinder coaxial with the rotation axis and consider the total torque exerted by the interior of the cylinder on its exterior. In collisionless disks, LBK's torque consists of two parts: a gravitational stress that is quadratic in gradients of the potential, and a "lorry transport" term that involves correlations between the radial and azimuthal velocities of stars crossing the cylinder. LBK's angular momentum flux is Eulerian—it describes the transfer of angular momentum from one spatial region to another.

Lagrangian approaches define the flux by the rate of transfer of angular momentum from one mass element to another. Lagrangian fluxes arise naturally via Noether's theorem from an action principle in which the dynamical variables follow the mass, but can also be defined directly.

For our stationary spiral waves, we define the angular momentum flux to be *the total torque exerted by the disk interior to a given streamline on the exterior.* No "lorry transport" term occurs because the gas does not cross streamlines, but the torque consists of two parts, a gravitational term and a pressure term. Appendix C shows that these contributions to the flux can be obtained conveniently by differentiating the internal and gravitational energies of the disk with respect to radial wavenumber α.‡ Thus the prescription for the flux is the same as in the tight-winding theory; however, the gravitational flux so obtained is strictly accurate only for exact logarithmic spirals in disks where $\Sigma_0(a) \propto a^{-3/2}$.

For as much of the stationary sequence as we have computed, the total angular momentum flux of the trailing self-similar spirals is always positive and increases monotonically with streamline eccentricity (Fig. 2).

‡ The total internal and gravitational energies of the disk are infinite, but we use the energies per unit logarithmic interval in semimajor axis, which are finite.

4 SUMMARY AND DISCUSSION

We have analysed stationary single-armed spiral density waves in disks whose rotation curve is strongly dominated by the potential of a pointlike central mass. Both linear and nonlinear waves have been considered. For tightly-wound waves, we have applied a phase-averaged variational formalism to derive the nonlinear dispersion relation, angular momentum density and flux, and characteristic velocities of these waves (Section 2 and Appendices A & B). The variational approach may be useful for other nonlinear dispersive waves of astrophysical interest (Whitham 1974). Unfortunately, when the waves are strongly nonlinear or when the disk is barely cool enough to permit their existence [eqs. (2) & (39)], the pitch angle is too large for the tightwinding approximation. Therefore we have constructed radially self-similar logarithmic spiral solutions without restriction on the pitch angle, but only for disks with azimuthally-averaged surface-density profiles scaling as $r^{-3/2}$ (Section 3). Where both are valid, the self-similar and tight-winding theories agree well. The most important lessons learned are as follows (unless otherwise noted, these statements apply to stationary, single-armed waves only):

(i) Even though self-gravity is essential, the Toomre Q parameter is irrelevant to these waves. This is because Q depends upon the central mass (via the orbital or epicyclic frequency), whereas the existence—and even the wavelength or pitch angle—of these waves is asymptotically independent of the mass ratio $M_{\rm central}/M_{\rm disk}$, since the waves consist of streamlines approximating keplerian ellipses. The appropriate dimensionless ratio of temperature to self-gravity is not Q^2 but rather σ^2 [eq. (2)], which is equivalent to the ratio of Jeans length to radius. A corollary is that the waves can propagate arbitrarily close to the central mass provided σ^2 remains below its critical value (39).

(ii) These waves are nonlinear extensions of the short branch of the WKB dispersion relation (at zero pattern speed and azimuthal wavenumber $m=1$); the long branch does not exist, at least not in isentropic disks with surface densities $\Sigma \propto r^{-3/2}$ (Section 3.1). Consequently, trailing waves propagate inward, even nonlinearly (Appendix B). Of course the long wave does exist at other arm multiplicities or retrograde pattern speeds.

(iii) If the azimuthally-averaged surface density profile is shallower (steeper) than $r^{-3/2}$, then ingoing (outgoing) waves become increasingly nonlinear as they propagate. This is a rather general result; it follows from the fact that the gravitational angular momentum flux scales as $G\Sigma^2 r^3$ at fixed pitch angle and eccentricity.

Regarding the first point above, it is always easier to have $\sigma^2 < 0.64$ in a keplerian disk than to have $Q \lesssim 1$ since

$$\sigma^2 = \frac{H}{2r} Q, \tag{54}$$

where $H \equiv v_s/\Omega$ is the disk thickness. Of course the single-armed waves we have studied are not locally unstable modes. On the other hand, it is interesting that the Jeans length $\lambda_J = 2\pi r \sigma^2$ is approximately the wavelength of fastest growth for *axisymmetric secular* instabilities (Lynden-Bell & Pringle 1974, Fridman & Polyachenko 1984). Viscous effects violate the conservation of vorticity, which tends to stabilize the disk at long wavelengths. Secular instability does not actually occur at large Q, however, unless the viscous torque is a constant or decreasing function of surface density (Schmit & Tscharnuter 1995).

With regard to the second point, it should be noted that although the waves we study belong formally to the short branch, their wavelengths are nevertheless long compared to the disk thickness when $Q \gg 1$. In fact their characteristic wavelength is $\lambda_J = \pi Q H$.

It is interesting that the minimum mass solar nebula, defined by augmenting the present planets with enough hydrogen and helium to reach solar abundance, has an $r^{-3/2}$ surface density profile. For example, Hayashi, Nakazawa, & Nakagawa 1985 quote

$$\Sigma(r) = 1.7 \times 10^3 \left(\frac{r}{1\ \mathrm{AU}}\right)^{-3/2} \mathrm{g\ cm^{-3}}, \qquad T(r) = 280 \left(\frac{L_*}{L_\odot}\right)^{1/4} \left(\frac{r}{1\ \mathrm{AU}}\right)^{-1/2}. \tag{55}$$

Very possibly this is only a coincidence. On the other hand, if density waves of the sort we have analysed were to dominate angular momentum transport, they might drive the surface density towards this power law: for a shallower profile, ingoing trailing waves would tend to steepen and dissipate at small radii, depositing negative angular momentum and driving accretion. The temperature profile above is shallower than expected for a thin optically thick disk warmed either by accretion or by reprocessing of solar radiation ($T \propto r^{-3/4}$); however, the $r^{-1/2}$ scaling seems required to match integrated infrared spectra (Adams, Lada, & Shu 1987) and may result from flaring of the disk (Kenyon & Hartmann 1987; Chiang & Goldreich 1997). For the profiles (55), the parameter σ^2 is independent of radius: $\sigma^2 \approx 1.5(L_*/L_\odot)^{1/4}$. This is comparable to the critical value 0.64. Under the same conditions, $Q \approx 70(r/\mathrm{AU})^{-1/4}$.

VLBI observations of extragalactic H_2O water masers have revealed molecular disks in the nuclei of NGC 4258 (Miyoshi et al. 1995) and NGC 1068 (Greenhill et al. 1996) on subparsec scales. In the former case, the disk appears to be very thin, though warped, and the rotation accurately keplerian, being dominated by a $3.6 \times 10^7\ M_\odot$ black hole. If one assumes that the observed Xray luminosity of the nucleus, $L_X = 4 \times 10^{40}$ erg s^{-1} (Makishima et al. 1994), is powered by *steady* accretion through this disk, then (cf. Neufeld & Maloney 1995)

$$\sigma^2 \approx 7 \times 10^{-3} \epsilon_{0.01} \alpha T_{300}^2 r_{0.1}^{3/2}, \qquad Q \approx 14. \epsilon_{0.01} \alpha T_{300}^{3/2} r_{0.1},$$

where $100\epsilon_{0.01}$ is the efficiency with which mass is converted to Xrays, $\alpha \lesssim 1$ is the usual viscosity parameter, T_{300} is the midplane temperature in units of 300 K (the minimum for masing), and $10 r_{0.1}$ pc is the distance from the black hole. Given the uncertainty in the first three parameters, it could easily be that $Q \lesssim 1$, but in any case the single-armed waves are surely permitted; the question is whether there is any way to excite them.

ACKNOWLEDGMENTS

We thank Scott Tremaine and Steve Lubow for helpful discussions, and the Isaac Newton Institute for Mathematical Sciences for its hospitality to JG while some of this work was carried out. This research was supported by NASA Astrophysical Theory Grant NAG5-2796.

REFERENCES

Abramowitz M., Stegun I. A., 1972, Handbook of Mathematical Functions, tenth printing (Washington, D.C.: U.S. Gov. Printing Ofc), Section 15.3.1
Adams F. C., Lada C. J., Shu F. H., 1987, ApJ, 312, 788
Adams F. C., Ruden S. P., Shu F. H., 1989, ApJ, 347, 959
Binney J. Tremaine S., 1987, Galactic Dynamics, Princeton Univ. Press, Princeton, NJ
Borderies N., Goldreich P., Tremaine, S., 1985, Icarus, 63, 406
—, 1986, Icarus, 68, 522
Brouwer D., Clemence G. M., 1961. Methods of Celestial Mechanics, Academic, New York
Chiang E. L., Goldreich, P., 1997, ApJ, 490, 368
Fridman, A. M., Polyachenko V. L., 1984, Physics of Gravitating Systems, vol. 2, Springer Verlag, New York
Goldreich P., Lynden-Bell D., 1965, MNRAS, 130, 125
Goldreich P., Tremaine S., 1979, ApJ, 233, 857
Goldreich P., Ward W. W., 1973, ApJ, 183, 1051
Greenhill L. J., Gwinn C. R., Antonucci, R., Barvainis, R. 1996, ApJ, 472, L21
Hayashi C., Nakazawa K., Nakagawa Y., 1985, in Black D. C., Matthews M. S., eds, Protostars and Planets II, Univ. of Arizona Press, Tucson, AZ, p. 1100
Heemskerk, M. H. M., Papaloizou, J. C., & Savonije, G. J., 1992, A&A, 260, 161
Kalnajs A., 1971, ApJ, 166, 275.
Kenyon S. J., Hartman L., 1997, ApJ, 323, 714
Lynden-Bell D., Kalnajs A., 1972, MNRAS, 157, 1 [LBK]
Lynden-Bell D., Ostriker J. P., 1967, MNRAS, 136, 293
Lynden-Bell D., Pringle J. E., 1974, MNRAS, 168, 603
Makishima K., et al., 1994, PASJ, 46, L77
Maoz E., 1995, ApJ, 455, L131
Miyoshi M., Moran J., Herrnstein J., Greenhill L., Nakai N., Diamond P., Inoue M., 1995, Nature, 373, 127
Neufeld D. A., Maloney P. R., 1995, ApJ, 447, L17
Safranov V. S., 1969, Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets, Nauka Press, Moscow
Schmit U., Tscharnuter W. M., 1995, Icarus, 115, 304
Shu F. H., Dones L., Lissauer J. J., Yuan C., Cuzzi J. N., 1985, ApJ, 299, 542
Shu F. H., Tremaine S., Adams F. C., Ruden S. P., 1990, ApJ, 358, 495
Shu F. H., Yuan C., Lissauer J. J., 1985, ApJ, 291, 356
Syer D., Tremaine S., 1996, MNRAS, 281, 925
Toomre A., 1964, ApJ, 139, 1217
Whitham G., 1974, Linear and Nonlinear Waves, Wiley, New York
Yuan C., Cassen P., 1985, Icarus, 64, 435
Yuan C., Cheng Y., 1989, ApJ, 340, 216
Yuan C., Cheng Y., 1991, ApJ, 376, 104

APPENDIX A: AVERAGES OVER ORBITAL PHASE FOR TIGHTLY-WOUND WAVES

After substitution from eq. (21), the phase average of the internal energy (16) becomes

$$\bar{W}_{\rm int} = \int 2\pi a da \frac{\Sigma c^2}{\gamma(\gamma-1)} \int_0^{2\pi} \frac{d\Psi}{2\pi} \left[(1 - q_0 \sin\Psi)^{-(\gamma-1)} - 1 \right] \quad = \int 2\pi a da \left[\frac{\Sigma c^2}{4} q_0^2 \mathcal{U}(q_0) \right]. \tag{A1}$$

The phase-averaged expression is independent of θ, integration over which has yielded the factor of 2π. The definition (24) of $\mathcal{U}(q_0)$ is closely related to an integral representation of the hypergeometric function (Abramowitz & Stegun 1972), whence

$$\mathcal{U}(q_0^2) = \frac{4}{\gamma(\gamma-1)q_0^2} \left[{}_2F_1\left(\frac{\gamma-1}{2}, \frac{\gamma}{2}, 1; q_0^2 \right) \; - 1 \right]. \tag{A2}$$

In the gravitational integral (17), we replace the limits on $\Delta a \equiv a_1 - a_2$ with $\pm\infty$ (a good approximation for tightly-wound waves) and eliminate θ in favor of Ψ using (18):

$$W_{\rm grav} = 2G \int a\Sigma^2(a) da \int_0^\infty d\Delta a \int_0^{2\pi} d\Psi \ln \left| 1 + \frac{ae\cos(\Psi + k\Delta a/2) - ae\cos(\Psi - k\Delta a/2)}{\Delta a} \right| \tag{A3}$$

The integration over Ψ can be cast in the form

$$\int_0^{2\pi} \frac{d\Psi}{2\pi} \ln|1 - R\sin\Psi| \;\; = \;\; \ln\left| \frac{1}{2} + \frac{1}{2}\sqrt{1-R^2} \right| \tag{A4}$$

where $R \equiv 2ae\sin(k\Delta a)/\Delta a$, or equivalently $R = q_0\sin(q_0 y)/y$ if $y \equiv \Delta a/2ae$. The integration over Δa can then be expressed in terms of the function $\mathcal{W}(q_0)$ defined in eq. (25). The gravitational energy is therefore

$$\bar{W}_{\rm grav} = \int 2\pi a da \left[-\frac{\pi G\Sigma^2}{2} ae|q_0|\mathcal{W}(q_0) \right]. \tag{A5}$$

We have not been able to relate $\mathcal{W}(q_0)$ to familiar special functions, so we evaluate it by expanding the the integrand of eq. (25) in powers of q_0^2 and integrating term by term.

APPENDIX B: NONLINEAR CHARACTERISTIC VELOCITIES

The characteristic velocities of a nonlinear wave train represent the speeds of propagation for modulations of the amplitude, wavelength, and frequency. Their derivation is simplified in our case by the special structure of the action (26)-(27) in the nearly stationary limit:

$$I \;\; = \;\; \int\int \left[\mathcal{L}_0(A,k,a) - \omega A \right] J(a) da dt \; + O(\omega^2), \tag{B1}$$

$$\mathcal{L}_0 \;\; \equiv \;\; -\frac{\pi G\Sigma_0 a}{2a\Omega} \left[\frac{1}{2} g e^2 - |q_0|\mathcal{W}(q_0) e + \sigma^2 q_0^2 \mathcal{U}(q_0) \right], \tag{B2}$$

where $A \equiv e^2/2$ and $J(a) \equiv 2\pi a^3 \Sigma_0(a)\Omega(a)$. The $O(\omega^2)$ terms do not not contribute to the characteristic velocities as $\omega \to 0$ and will be neglected. Variation of I with respect to A yields the dispersion relation

$$\omega(A,k,a) = \frac{\partial \mathcal{L}_0}{\partial A}, \tag{B3}$$

and variation with respect to the phase function Ψ yields the conservation of wave action (cf. eq. [11])

$$\frac{\partial}{\partial t}(JA) + \frac{\partial}{\partial a}\left(J \frac{\partial \mathcal{L}_0}{\partial k} \right) = 0. \tag{B4}$$

On the other hand, because partial derivatives of Ψ commute, it follows from eqs. (8) that

$$\frac{\partial k}{\partial t} + \frac{\partial \omega}{\partial a} = 0. \tag{B5}$$

The derivatives with respect to a in equations (B4) and (B5) include the indirect dependence of $\mathcal{L}_0$ and ω on a via their dependence on k and A. Substituting from eq. (B3) and expanding derivatives, we have

$$\begin{aligned} \frac{\partial k}{\partial t} \; + \; \left(\frac{\partial^2 \mathcal{L}_0}{\partial A \partial k} \right) \frac{\partial k}{\partial a} \; + \; \left(\frac{\partial^2 \mathcal{L}_0}{\partial A^2} \right) \frac{\partial A}{\partial a} \;\; &= \;\; -\frac{\partial \mathcal{L}_0}{\partial a} \\ \frac{\partial A}{\partial t} \; + \; \left(\frac{\partial^2 \mathcal{L}_0}{\partial k^2} \right) \frac{\partial k}{\partial a} \; + \; \left(\frac{\partial^2 \mathcal{L}_0}{\partial k \partial A} \right) \frac{\partial A}{\partial a} \;\; &= \;\; -\frac{\partial^2 \mathcal{L}_0}{\partial a \partial k} - \frac{\partial \mathcal{L}_0}{\partial a}\frac{\partial \ln J}{\partial a}, \end{aligned} \tag{B6}$$

where now $\partial \mathcal{L}_0/\partial a$ involves only the explicit dependence of $\mathcal{L}_0(A,k,a)$ on its third argument.

Equations (B6) form a system of first-order quasilinear partial differential equations for (k, A) in the (a, t) plane. Such systems can be solved by integrating along characteristic curves $da/dt = V(k, A, a, t)$. The characteristic velocities V are eigenvalues of the matrix of coefficients of $\left(\frac{\partial k}{\partial a}, \frac{\partial A}{\partial a}\right)$:

$$V_1, V_2 = \frac{\partial^2 \mathcal{L}_0}{\partial k \partial A} \pm \sqrt{\frac{\partial^2 \mathcal{L}_0}{\partial A^2}\frac{\partial^2 \mathcal{L}_0}{\partial k^2}} = \frac{\partial \omega}{\partial k} \pm \sqrt{\frac{\partial \omega}{\partial A}\frac{\partial^2 \mathcal{L}_0}{\partial k^2}}\,. \tag{B7}$$

The reduced lagrangian (B2) has a Taylor series in A beginning at $O(A)$, and ω has a Taylor series beginning at $O(1)$. Therefore as $A \to 0$, the characteristic velocities merge and can be identified with the linear group velocity, but at finite amplitude they are distinguished by corrections of order $A^{1/2} = O(e)$. Should the argument of the surd in (B7) be negative, the characteristic velocities are complex; this may signal a tendency of the nonlinear wave train to break up into a series of solitary waves (Whitham 1974). The required derivatives of $\mathcal{L}_0$ are

$$\begin{aligned}
\frac{\partial^2 \mathcal{L}_0}{\partial k \partial A} = \frac{\partial \omega}{\partial k} &= \frac{\pi G \Sigma}{2\Omega}\left[-\sigma^2 ka\left(q_0^2 \mathcal{U}'' + 5q_0\mathcal{U}' + 4\mathcal{U}\right) + \left(q_0^2\mathcal{W}'' + 4q_0\mathcal{W}' + 2\mathcal{W}\right)\operatorname{sign}(k)\right],\\
\frac{\partial^2 \mathcal{L}_0}{\partial A^2} = \frac{\partial \omega}{\partial A} &= \frac{\pi G \Sigma}{2\Omega}\frac{k^2 a}{e^2}\left[-\sigma^2\left(q_0^2\mathcal{U}'' + 3q_0\mathcal{U}'\right) + e\left(q_0\mathcal{W}'' + 3\mathcal{W}'\right)\operatorname{sign}(k)\right],\\
\frac{\partial^2 \mathcal{L}_0}{\partial k^2} &= \frac{\pi G \Sigma}{2\Omega} a e^2\left[-\sigma^2\left(q_0^2\mathcal{U}'' + 4q_0\mathcal{U}' + 2\mathcal{U}\right) + e\left(q_0\mathcal{W}'' + 2\mathcal{W}'\right)\operatorname{sign}(k)\right].
\end{aligned} \tag{B8}$$

The characteristic velocities are shown for representative values of σ^2 in Figure 3. We have plotted results for the short branch only, since the the tight-winding approximation is not applicable to the long branch. In all cases that we have examined, both chararacteristic velocities are real and negative for tightly wrapped trailing waves ($ka \gg 1$, $k > 0$.)

APPENDIX C: ANGULAR MOMENTUM FLUX OF SELF-SIMILAR SPIRALS

We relate the angular momentum flux of nonlinear logarithmic spirals to derivatives of the gravitational and internal energies with respect to pitch angle. The flux proves to be constant in isentropic disks with $\Sigma_0(a) \propto a^{-3/2}$ and adiabatic index $\gamma = 4/3$, so that no net torque is exerted on streamlines.

Consider first the gravitational interaction energy between two keplerian elliptical rings of the same eccentricity e, semimajor axes (a_1, a_2), periapse angles ϕ_1, ϕ_2, and masses (m_1, m_2). After integration over the arc lengths, the energy can be put into the form [cf. eqs (44) & (46)]

$$W_{12} = -\left(\frac{G m_1 m_2}{\sqrt{a_1 a_2}}\right) K\left(\ln\frac{a_2}{a_1},\ \phi_2 - \phi_1\right). \tag{C1}$$

Because we have factored out $(a_1 a_2)^{-1/2}$, the dimensionless kernel $K(\xi, \phi)$ is independent of the absolute dimensions of the rings, and it is symmetric in the sign of both arguments. The torque exerted by ring 2 on ring 1 is, with $x \equiv \ln a$,

$$\Gamma_{12} = -\frac{\partial W_{12}}{\partial \phi_1} = -\frac{G m_1 m_2}{\sqrt{a_1 a_2}} K_\phi(x_2 - x_1, \phi_2 - \phi_1), \tag{C2}$$

where $K_\phi(\xi, \phi) \equiv \partial K(\xi, \phi)/\partial\phi$. In a continuous elliptical disk, let

$$\mu \equiv 2\pi a^{3/2}\Sigma_0(a), \tag{C3}$$

so that the mass between streamlines a and $a + da$ is $a^{1/2}\mu dx$, and μ is constant since $\Sigma_0 \propto a^{-3/2}$. The gravitational energy per logarithmic interval in semimajor axis is therefore

$$W(x_1, \phi_1) = -\frac{G\mu^2}{2}\int_{-\infty}^{\infty} K(x_2 - x_1, \phi_2 - \phi_1)dx_2, \tag{C4}$$

and the total torque per unit x is

$$\frac{\partial W}{\partial \phi_1}(x_1, \phi_1) = -\frac{G\mu^2}{2}\int_{-\infty}^{\infty} K_\phi(x_2 - x_1, \phi_2 - \phi_1)dx_2. \tag{C5}$$

In a logarithmic spiral, $\phi_2 - \phi_1 = -\alpha \cdot (x_2 - x_1)$. Then the energy (C4) is independent of x_1. However, the energy does depend upon the pitch angle:

$$\frac{\partial W}{\partial \alpha} = G\mu^2 \int_0^\infty \xi K_\phi(\xi, -\alpha\xi) d\xi. \tag{C6}$$

The total torque exerted by all rings $x_2 < x$ on all rings $x_1 > x$ does not vanish. Referring to eq. (C2), and noting the antisymmetry of K_ϕ in its second argument, this torque is

$$\begin{aligned} \Gamma_{\rm grav}(x,\alpha) &= +G\mu^2 \int_x^\infty dx_1 \int_{-\infty}^x dx_2 K_\phi[x_1 - x_2, -\alpha(x_1 - x_2)] \\ &= G\mu^2 \int_0^\infty d\xi \int_{x-\xi/2}^{x+\xi/2} d\eta K_\phi(\xi, -\alpha\xi) \;=\; G\mu^2 \int_0^\infty \xi K_\phi(\xi, -\alpha\xi) d\xi \;=\; \frac{\partial W}{\partial \alpha}. \end{aligned} \tag{C7}$$

This establishes the desired relationship between the gravitational contribution to the angular momentum flux and the derivative of the gravitational energy with respect to pitch angle. Since $\partial \Gamma_{\rm grav}/\partial x = 0$, there is no net torque on streamlines.

Now consider the pressure term. For a polytropic gas, the internal energy per unit mass is

$$\frac{p}{\gamma - 1} = \frac{p_0}{\gamma - 1}\left(\frac{V}{V_0}\right)^{-\gamma},$$

where p is the pressure, $V = 1/\Sigma$ is the specific volume, and the subscript "0" denotes a reference state on the gas adiabat. Integrating around a streamline, and noting the area element $a^2 \mathcal{J} dE dx$, we have for the internal energy per logarithmic interval of semimajor axis,

$$U(a,\alpha) = \frac{p_0(a)a^2}{\gamma - 1} \int_0^{2\pi} \mathcal{J}^{1-\gamma} dE, \tag{C8}$$

where E is the eccentric anomaly and $\mathcal{J}$ is the Jacobian (48). The derivative with respect to α is

$$\frac{\partial U}{\partial \alpha} = p_0(a)a^2 \int_0^{2\pi} e \sin E \mathcal{J}^{-\gamma} dE. \tag{C9}$$

The angular momentum flux due to pressure is the integral of $r\boldsymbol{e}_\theta \cdot \boldsymbol{T} \cdot \boldsymbol{n} = rp\boldsymbol{e}_\theta \cdot \boldsymbol{n}$ along the arc length of the streamline, where $\boldsymbol{T}$ is the pressure tensor, $\boldsymbol{e}_\theta$ a unit azimuthal vector, and $\boldsymbol{n}$ the outward unit normal to the streamline. If $a = a(r,\theta)$ then $\boldsymbol{n} = |\boldsymbol{\nabla} a|^{-1}\boldsymbol{\nabla} a$, and the area element can be written $|\boldsymbol{\nabla} a|^{-1} ds da = \mathcal{J} a da dE$, ds being the element of arc length. Hence,

$$\Gamma_{\rm press} = \int_0^{2\pi} p(a,E) a \mathcal{J} \left(\frac{\partial a}{\partial \theta}\right)_r dE. \tag{C10}$$

Now $\gamma p(a,E) = \Sigma_0(a) c_{s,0}^2(a) \mathcal{J}^{-\gamma}$, and equations (44) yield $(\partial a/\partial\theta)_r = -\mathcal{J}^{-1} ae \sin E$. With these substitutions, we have

$$\Gamma_{\rm press} = \frac{\partial U}{\partial \alpha}, \tag{C11}$$

as claimed. Since the prefactor $p_0^2(a)a^2$ is independent of a in our self-similar spiral disks, $\Gamma_{\rm press}$ is also independent of a, so that the net pressure torque on a streamline vanishes.

This paper has been produced using the Royal Astronomical Society/Blackwell Science LaTeX style file.